



1

SEQUENCE LISTING

#37  
RECEIVED  
NOV 23 2001  
TECH CENTER 1600/2900

(1) GENERAL INFORMATION:

- (i) APPLICANT: ANDERSON, DARRELL R.  
HANNA, NABIL  
LEONARD, JOHN E.  
NEWMAN, ROLAND A.  
REFF, MITCHELL E.  
RASTETTER, WILLIAM H.
- (ii) TITLE OF INVENTION: THERAPEUTIC APPLICATION OF CHIMERIC AND  
RADIOLABELED ANTIBODIES TO HUMAN B LYMPHOCYTE RESTRICTED  
DIFFERENTIATION ANTIGEN FOR THE TREATMENT OF B CELL  
LYMPHOMA
- (iii) NUMBER OF SEQUENCES: 11
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: PILLSBURY WINTHROP
  - (B) STREET: 1100 New York Avenue, N.W., Ninth FL.
  - (C) CITY: Washington
  - (D) STATE: DC
  - (E) COUNTRY: USA
  - (F) ZIP: 20005
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/921,060
  - (B) FILING DATE: 29-AUG-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/149,099
  - (B) FILING DATE: 03-NOV-1993
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 07/978,891
  - (B) FILING DATE: 13-NOV-1992
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Teskin, Robin L.
  - (B) REGISTRATION NUMBER: 35,030
  - (C) REFERENCE/DOCKET NUMBER: 037003-0275463
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 202-861-3000
  - (B) TELEFAX: 202-822-0944

(2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGGAGCTTGG ATCGATCCTC TATGGTT

27

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8540 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GACGTCGCGG CCGCTCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG 60  
 AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGCATGGGGC 120  
 GGAGAATGGG CGGAAGTGGG CGGAGTTAGG GCGGGGATGG GCGGAGTTAG GGGCGGGACT 180  
 ATGGTTGCTG ACTAATTGAG ATGCATGCTT TGCATACTTC TGCCTGCTGG GGAGCCTGGG 240  
 GACTTTCCAC ACCTGGTTGC TGACTAATTG AGATGCATGC TTTGCATACT TCTGCCTGCT 300  
 GGGGAGCCTG GGGACTTTCC ACACCCTAAC TGACACACAT TCCACAGAAAT TAATTCCCCT 360  
 AGTTATTAAT AGTAATCAAT TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC 420  
 GTTACATAAC TTACGGTAAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATG 480  
 ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA TTGACGTCAA 540  
 TGGGTGGACT ATTTACGGTA AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA 600  
 AGTACGCCCC CTATTGACGT CAATGACGGT AAATGGCCCC CCTGGCATTG TGCCAGTAC 660  
 ATGACCTTAT GGGACTTTCC TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC 720  
 ATGGTGATGC GGTTTTGGCA GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA 780  
 TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA AAATCAACGG 840  
 GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA 900  
 CGGTGGGAGG TCTATATAAG CAGAGCTGGG TACGTGAACC GTCAGATCGC CTGGAGACGC 960  
 CATCACAGAT CTCTCACCAT GAGGGTCCCC GCTCAGCTCC TGGGGCTCCT GCTGCTCTGG 1020  
 CTCCCAGGTG CACGATGTGA TGGTACCAAG GTGGAATCA AACGTACGGT GGCTGCACCA 1080  
 TCTGTCTTCA TCTTCCCGCC ATCTGATGAG CAGTTGAAAT CTGGAAGTGC CTCTGTTGTG 1140  
 TGCCTGCTGA ATAACCTTCTA TCCCAGAGAG GCCAAAGTAC AGTGGAAGGT GGATAACGCC 1200  
 CTCCAATCGG GTAACCTCCA GGAGAGTGTC ACAGAGCAGG ACAGCAAGGA CAGCACCTAC 1260  
 AGCCTCAGCA GCACCCTGAC GCTGAGCAAA GCAGACTACG AGAAACACAA AGTCTACGCC 1320  
 TGCGAAGTCA CCCATCAGGG CCTGAGCTCG CCCGTCACAA AGAGCTTCAA CAGGGGAGAG 1380  
 TGTTGAATTC AGATCCGTTA ACGGTTACCA ACTACCTAGA CTGGATTTCG GACAACATGC 1440  
 GGCCGTGATA TCTACGTATG ATCAGCCTCG ACTGTGCCTT CTAGTTGCCA GCCATCTGTT 1500  
 GTTTGCCCTT CCCCCGTGCC TTCCTTGACC CTGGAAGGTG CCACTCCAC TGTCCTTTCC 1560  
 TAATAAAATG AGGAAATTGC ATCGCATTGT CTGAGTAGGT GTCATTCTAT TCTGGGGGGT 1620  
 GGGGTGGGGC AGGACAGCAA GGGGGAGGAT TGGGAAGACA ATAGCAGGCA TGCTGGGGAT 1680

GCGGTGGGCT	CTATGGAACC	AGCTGGGGCT	CGACAGCTAT	GCCAAGTACG	CCCCCTATTG	1740
ACGTCAATGA	CGGTAAATGG	CCCCGCTGGC	ATTATGCCCA	GTACATGACC	TTATGGGACT	1800
TTCCTACTTG	GCAGTACATC	TACGTATTAG	TCATCGCTAT	TACCATGGTG	ATGCGGTTTT	1860
GGCAGTACAT	CAATGGGCGT	GGATAGCGGT	TTGACTCACG	GGGATTTCCA	AGTCTCCACC	1920
CCATTGACGT	CAATGGGAGT	TTGTTTTGGC	ACCAAAATCA	ACGGGACTTT	CCAAAATGTC	1980
GTAACAATC	CGCCCCATTG	ACGCAAATGG	GCGGTAGGCG	TGTACGGTGG	GAGGTCTATA	2040
TAAGCAGAGC	TGGGTACGTC	CTCACATTCA	GTGATCAGCA	CTGAACACAG	ACCCGTCGAC	2100
ATGGGTTGGA	GCCTCATCTT	GCTCTTCCTT	GTGCTGTTG	CTACGCGTGT	CGCTAGCACC	2160
AAGGGCCCCAT	CGGTCTTCCC	CCTGGCACCC	TCCTCCAAGA	GCACCTCTGG	GGGCACAGCG	2220
GCCCTGGGCT	GCCTGGTCAA	GGACTACTTC	CCCGAACCGG	TGACGGTGTG	GTGGAATCTA	2280
GGCGCCCTGA	CCAGCGGCGT	GCACACCTTC	CCGGCTGTCC	TACAGTCCTC	AGGACTCTAC	2340
TCCCTCAGCA	CGGTGGTGAC	CGTGCCCTCC	AGCAGCTTGG	GCACCCAGAC	CTACATCTGC	2400
AACGTGAATC	ACAAGCCCAG	CAACACCAAG	GTGGACAAGA	AAGCAGAGCC	CAAATCTTGT	2460
GACAAAATC	ACACATGCCC	ACCGTGCCCCA	GCACCTGAAC	TCCTGGGGGG	ACCGTCAGTC	2520
TTCCTCTTCC	CCCCAAAACC	CAAGGACACC	CTCATGATCT	CCCGGACCCC	TGAGGTCACA	2580
TGCGTGGTGG	TGGACGTGAG	CCACGAAGAC	CCTGAGGTCA	AGTTCAACTG	GTACGTGGAC	2640
GGCGTGGAGG	TGCATAATGC	CAAGACAAAG	CCGCGGGAGG	AGCAGTACAA	CAGCACGTAC	2700
CGTGTGGTCA	GCGTCCTCAC	CGTCCCTGCAC	CAGGACTGGC	TGAATGGCAA	GGACTACAAG	2760
TGCAAGGTCT	CCAACAAAGC	CCTCCCAGCC	CCCATCCGAG	AAACCATCTC	CAAAGCCAAA	2820
GGGCAGCCCC	GAGAACCCACA	GGTGTAACAC	CTGCCCCAT	CCCGGGATGA	GCTGACCAGG	2880
AACCAGGTCA	GCCTGACCTG	CCTGGTCAAA	GGCTTCTATC	CCAGCGACAT	CGCCGTGGAG	2940
TGGGAGAGCA	ATGGGCAGCC	GGAGAACAAC	TACAAGACCA	CGCCTCCCGT	GCTGGACTCC	3000
GACGGCTCCT	TCTTCTCTTA	CAGCAAGCTC	ACCGTGGACA	AGAGCAGGTG	GCAGCAGGGG	3060
AACGTCTTCT	CATGCTCCGT	GATGCATGAG	GCTCTGCACA	ACCACTACAC	GCAGAAGAGC	3120
CTCTCCCTGT	CTCCGGGTAA	ATGAGGATCC	GTTAACGGTT	ACCAACTACC	TAGACTGGAT	3180
TCGTGACAAC	ATGCGGCCGT	GATATCTACG	TATGATCAGC	CTCGACTGTG	CCTTCTAGTT	3240
GCCAGCCATC	TGTTGTTTTG	CCCTCCCCCG	TGCCTTCCTT	GACCCTGGAA	GGTGCCACTC	3300
CCACTGTCCT	TTCCTAATAA	AATGAGGAAA	TTGCATCGCA	TTGTCTGAGT	AGGTGTCATT	3360
CTATTCTGGG	GGGTGGGGTG	GGGCAGGACA	GCAAGGGGGA	GGATTGGGAA	GACAATAGCA	3420
GGCATGCTGG	GGATGCGGTG	GGCTCTATGG	AACCAGCTGG	GGCTCGACAG	CGCTGGATCT	3480
CCCGATCCCC	AGCTTTGCTT	CTCAATTTCT	TATTTGCATA	ATGAGAAAAA	AAGGAAAAAT	3540
AATTTTAACA	CCAATTCAGT	AGTTGATTGA	GCAAATGCGT	TGCCAAAAAG	GATGCTTTAG	3600
AGACAGTGTT	CTCTGCACAG	ATAAGGACAA	ACATTATTCA	GAGGGAGTAC	CCAGAGCTGA	3660
GACTCCTAAG	CCAGTGAGTG	GCACAGCATT	CTAGGGAGAA	ATATGCTTGT	CATCACCGAA	3720
GCCTGATTCC	GTAGAGCCAC	ACCTTGGTAA	GGGCCAATCT	GCTCACACAG	GATAGAGAGG	3780
GCAGGAGCCA	GGGCAGAGCA	TATAAGGTGA	GGTAGGATCA	GTTGCTCCTC	ACATTTGCTT	3840
CTGACATAGT	TGTGTTGGGA	GCTTGGATAG	CTTGGACAGC	TCAGGGCTGC	GATTTCCGCG	3900
CAAACCTGAC	GGCAATCCTA	GCGTGAAGGC	TGGTAGGATT	TTATCCCCGC	TGCCATCATG	3960
GTTTCGACCAT	TGAATGTCAT	CGTCGCCGTG	TCCCCAAAATA	TGGGGATTGG	CAAGAACGGA	4020
GACCTACCCT	GGCCTCCGCT	CAGGAACGAG	TTCAAGTACT	TCCAAAGAAT	GACCACAACC	4080
TCTTCAGTGG	AAGGTAAACA	GAATCTGGTG	ATTATGGGTA	GGAAAACCTG	GTTCTCCATT	4140
CCTGAGAACA	ATCGACCTTT	AAAGGACAGA	ATTAATATAG	TTCTCAGTAG	AGAACTCAAA	4200
GAACCACCAC	GAGGAGCTCA	TTTTCTTGCC	AAAAGTTTGG	ATGATGCCTT	AAGACTTATT	4260
GAACAACCGG	AATTGGCAAG	TAAAGTAGAC	ATGGTTTGGG	TAGTCGGAGG	CAGTTCGTGT	4320
TACCAGGAAG	CCATGAATCA	ACCAGGCCAC	CTTAGACTCT	TTGTGACAAG	GATCATGCAG	4380
GAATTTGAAA	GTGACACGTT	TTTCCCAGAA	ATTGATTTGG	GGAAATATAA	ACTTCTCCCA	4440
GAATACCCAG	GCGTCCTCTC	TGAGGTCCAG	GAGGAAAAAG	GCATCAAGTA	TAAGTTTGAA	4500
GTCTACGAGA	AGAAAGACTA	ACAGGAAGAT	GCTTTCAAGT	TCTCTGCTCC	CCTCCTAAAG	4560
TCATGCATTT	TTATAAGACC	ATGGGACTTT	TGCTGGCTTT	AGATCAGCCT	CGACTGTGCC	4620
TTCTAGTTGC	CAGCCATCTG	TTGTTTGCCC	CTCCCCCGTG	CCTTCCTTGA	CCCTGGAAGG	4680
TGCCACTCCC	ACTGTCCTTT	CCTAATAAAA	TGAGGAAATT	GCATCGCATT	GTCTGAGTAG	4740
GTGTCAATTCT	ATTCTGGGGG	GTGGGGTGGG	GCAGGACAGC	AAGGGGGAGG	ATTGGGAAGA	4800
CAATAGCAGG	CATGCTGGGG	ATGCGGTGGG	CTCTATGGAA	CCAGCTGGGG	CTCGAGCTAC	4860
TAGCTTTGCT	TCTCAATTTT	TTATTTGCAT	AATGAGAAAA	AAAGGAAAAAT	TAATTTTAAC	4920
ACCAATTCAG	TAGTTGATTG	AGCAAATGCG	TTGCCAAAAA	GGATGCTTTA	GAGACAGTGT	4980
TCTCTGCACA	GATAAGGACA	AACATTATTC	AGAGGGAGTA	CCCAGAGCTG	AGACTCCTAA	5040
GCCAGTGAGT	GGCACAGCAT	TCTAGGGAGA	AATATGCTTG	TCATCACCAG	AGCCTGATTC	5100
CGTAGAGCCA	CACCTTGCGT	AGGGCCAATC	TGCTCACACA	GGATAGAGAG	GGCAGGAGCC	5160

AGGGCAGAGC	ATATAAGGTG	AGGTAGGATC	AGTTGCTCCT	CACATTTGCT	TCTGACATAG	5220
TTGTGTTGGG	AGCTTGGATC	GATCCTCTAT	GGTTGAACAA	GATGGATTGC	ACGCAGGTTC	5280
TCCGGCCGCT	TGGGTGGAGA	GGCTATTTCG	CTATGACTGG	GCACAACAGA	CAATCGGCTG	5340
CTCTGATGCC	GCCGTGTTCC	GGCTGTCAGC	GCAGGGGCGC	CCGGTTCTTT	TTGTCAAGAC	5400
CGACCTGTCC	GGTGCCCTGA	ATGAACTGCA	GGACGAGGCA	GCGCGGCTAT	CGTGGCTGGC	5460
CACGACGGGC	GTTCTTGCG	CAGCTGTGCT	CGACGTTGTC	ACTGAAGCGG	GAAGGGACTG	5520
GCTGCTATTG	GGCGAAGTGC	CGGGGCAGGA	TCTCCTGTCA	TCTCACCTTG	CTCCTGCCGA	5580
GAAAGTATCC	ATCATGGCTG	ATGCAATGCG	GCGGCTGCAT	ACGCTTGATC	CGGTACCTG	5640
CCCATTCGAC	CACCAAGCGA	AACATCGCAT	CGAGCGAGCA	CGTACTCGGA	TGGAAGCCGG	5700
TCTTGTCGAT	CAGGATGATC	TGGACGAAGA	GCATCAGGGG	CTCGCGCCAG	CCGAACTGTT	5760
CGCCAGGCTC	AAGGCGCGCA	TGCCCCGACG	CGAGGATCTC	GTCGTGACCC	ATGGCGATGC	5820
CTGCTTGCCG	AATATCATGG	TGGAATAATG	CCGCTTTTCT	GGATTCATCG	ACTGTGGCCG	5880
GCTGGGTGTG	GCGGACCGCT	ATCAGGACAT	AGCGTTGGCT	ACCCGTGATA	TTGCTGAAGA	5940
GCTTGGCGGC	GAATGGGCTG	ACCGCTTCCT	CGTGCTTTAC	GGTATCGCCG	CTCCCGATTC	6000
GCAGCGCATC	GCCTTCTATC	GCCTTCTTGA	CGAGTTCTTC	TGAGCGGGAC	TCTGGGGTTC	6060
GAAATGACCG	ACCAAGCGAC	GCCCAACCTG	CCATCACGAG	ATTTGATTC	CACCGCCGCC	6120
TTCTATGAAA	GGTTGGGCTT	CGGAATCGTT	TTCCGGGACG	CCGGCTGGAT	GATCCTCCAG	6180
CGCGGGGATC	TCATGCTGGA	GTTCTTCCGC	CACCCCAACT	TGTTTATTGC	AGCTTATAAT	6240
GGTTACAAAT	AAAGCAATAG	CATCACAAT	TTCACAAATA	AAGCATTMTT	TTCACTGCAT	6300
TCTAGTTGTG	GTTTGTCCAA	ACTCATCAAT	CTATCTTATC	ATGTCTGGAT	CGCGGCCGCG	6360
ATCCCGTCGA	GAGCTTGGCG	TAATCATGGT	CATAGCTGTT	TCCTGTGTGA	AATTGTTATC	6420
CGCTCACAA	TCCACACAAC	ATACGAGCCG	GAAGCATAAA	GTGTAAAGCC	TGGGGTGCCT	6480
AATGAGTGAG	CTAACTCACA	TTAATTGCGT	TGCGCTCACT	GCCCGCTTTC	CAGTCGGGAA	6540
ACCTGTCGTG	CCAGCTGCAT	TAATGAATCG	GCCAACGCGC	GGGGAGAGGC	GGTTTGCGTA	6600
TTGGGCGCTC	TTCCGCTTCC	TCGCTCACTG	ACTCGCTGCG	CTCGGTGCTT	CGGCTGCGGC	6660
GAGCGGTATC	AGCTCACTCA	AAGGCGGTAA	TACGGTTATC	CACAGAATCA	GGGGATAACG	6720
CAGGAAAGAA	CATGTGAGCA	AAAGGCCAGC	AAAAGGCCAG	GAACCGTAAA	AAGGCCGCGT	6780
TGCTGGCGTT	TTTCCATAGG	CTCCGCCCCC	CTGACGAGCA	TCACAAAAAT	CGACCTCAA	6840
GTCAGAGGTG	GCGAAACCCG	ACAGGACTAT	AAAGATACCA	GGCGTTTCCC	CCTGGAAGCT	6900
CCCTCGTGCG	CTCTCCTGTT	CCGACCCTGC	CGCTTACCGG	ATACCTGTCC	GCCTTTCTCC	6960
CTTCGGGAAG	CGTGGCGCTT	TCTCAATGCT	CACGCTGTAG	GTATCTCAGT	TCGGTGTAGG	7020
TCGTTGCTC	CAAGCTGGGC	TGTGTGCACG	AACCCCCCGT	TCAGCCCCGAC	CGCTGCGCCT	7080
TATCCGGTAA	CTATCGTCTT	GAGTCCAACC	CGGTAAGACA	CGACTTATCG	CCACTGGCAG	7140
CAGCCACTGG	TAACAGGATT	AGCAGAGCGA	GGTATGTAGG	CGGTGCTACA	GAGTTCCTGA	7200
AGTGGTGGCC	TAACACGGC	TACACTAGAA	GGACAGTATT	TGGTATCTGC	GCTCTGCTGA	7260
AGCCAGTTAC	CTTCGGAAAA	AGAGTTGGTA	GCTCTTGATC	CGGCAAAACA	ACCACCGCTG	7320
GTAGCGGTGG	TTTTTTTGT	TGCAAGCAGC	AGATTACGCG	CAGAAAAAAA	GGATCTCAAG	7380
AAGATCCTTT	GATCTTTTCT	ACGGGGTCTG	ACGCTCAGTG	GAACGAAAAAC	TCACGTAAAG	7440
GGATTTTGGT	CATGAGATTA	TCAAAAAGGA	TCTTCACTTA	GATCCTTTTA	AATTAAAAAT	7500
GAAGTTTAA	ATCAATCTAA	AGTATATATG	AGTAAACTTG	GTCTGACAGT	TACCAATGCT	7560
TAATCAGTGA	GGCACCTATC	TCAGCGATCT	GTCTATTTTC	TTTATCCATA	GTTGCCTGAC	7620
TCCCGCTCGT	GATGATAACT	ACGATACGGG	AGGGCTTACC	ATCTGGCCCC	AGTGCTGCAA	7680
TGATACCGCG	AGACCCACGC	TCACCGGCTC	CAGATTTATC	AGCAATAAAC	CAGCCAGCCG	7740
GAAGGGCCGA	GCGCAGAAGT	GGTCCTGCAA	CTTTATCCGC	CTCCATCCAG	TCTATTAATT	7800
GTTGCCGGGA	AGCTAGAGTA	AGTAGTTCGC	CAGTTAATAG	TTTGCGCAAC	GTTGTTGCCA	7860
TTGCTACAGG	CATCGTGGTG	TCACGCTCGT	CGTTTGGTAT	GGCTTCATTC	AGCTCCGGTT	7920
CCCAACGATC	AAGGCGAGTT	ACATGATCCC	CCATGTTGTG	CAAAAAAGCG	GTTAGCTCCT	7980
TCGGTCCTCC	GATCGTTGTC	AGAAGTAAGT	TGGCCGCGAGT	GTTATCACTC	ATGGTTATGG	8040
CAGCACTGCA	TAATTCTCTT	ACTGTCATGC	CATCCGTAAG	ATGCTTTTCT	GTGACTGGTG	8100
AGTACTCAAC	CAAGTCATTC	TGAGAATAGT	GTATGCGGCG	ACCGAGTTGC	TCTTGCCCGG	8160
CGTCAATACG	GGATAATACC	GCGCCACATA	GCAGAACTTT	AAAAGTGCTC	ATCATTGGAA	8220
AACGTTCTTC	GGGGCGAAAA	CTCTCAAGGA	TCTTACCGCT	GTTGAGATCC	AGTTCGATGT	8280
AACCCACTCG	TGCACCCAAC	TGATCTTCAG	CATCTTTTAC	TTTACCAGC	GTTTCTGGGT	8340
GAGCAAAAAAC	AGGAAGGCAA	AATGCCGCAA	AAAAGGGAAT	AAGGGCGACA	CGGAAATGTT	8400
GAATACTCAT	ACTCTTCTTT	TTTCAATATT	ATTGAAGCAT	TTATCAGGGT	TATTGTCTCA	8460
TGAGCGGATA	CATATTTGAA	TGTATTTAGA	AAAAATAACA	AATAGGGGTT	CCGCGCACAT	8520
TTCCCCGAAA	AGTGCCACCT					8540

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9209 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GACGTCGCGG	CCGCTCTAGG	CCTCCAAAAA	AGCCTCCTCA	CTACTTCTGG	AATAGCTCAG	60
AGGCCGAGGC	GGCCTCGGCC	TCTGCATAAA	TAAAAAAAT	TAGTCAGCCA	TGCATGGGGC	120
GGAGAAATGGG	CGGAACTGGG	CGGAGTTAGG	GGCGGGATGG	GCGGAGTTAG	GGGCGGGACT	180
ATGGTTGCTG	ACTAATTGAG	ATGCATGCTT	TGCATACTTC	TGCCTGCTGG	GGAGCCTGGG	240
GACTTTCCAC	ACCTGGTTGC	TGACTAATTG	AGATGCATGC	TTTGCATACT	TCTGCCTGCT	300
GGGGAGCCTG	GGGACTTTCC	ACACCCTAAC	TGACACACAT	TCCACAGAAT	TAATTCCCCT	360
AGTTATTAAT	AGTAATCAAT	TACGGGGTCA	TTAGTTCATA	GCCCATATAT	GGAGTTCGCG	420
GTTACATAAC	TTACGGTAAA	TGGCCCGCCT	GGCTGACCGC	CCAACGACCC	CCGCCCATTG	480
ACGTCAATAA	TGACGTATGT	TCCCATAGTA	ACGCCAATAG	GGACTTTCCA	TTGACGTCAA	540
TGGGTGGACT	ATTTACGGTA	AACTGCCCCAC	TTGGCAGTAC	ATCAAGTGTA	TCATATGCCA	600
AGTACGCCCC	CTATTGACGT	CAATGACGGT	AAATGGCCCCG	CCTGGCATTG	TGCCCAGTAC	660
ATGACCTTAT	GGGACTTTCC	TACTTGGCAG	TACATCTACG	TATTAGTCAT	CGCTATTACC	720
ATGGTGATGC	GGTTTTGGCA	GTACATCAAT	GGGCGTGGAT	ACCGGTTTGA	CTCACGCGGA	780
TTTCCAAGTC	TCCACCCCAT	TGACGTCAAT	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	840
GACTTTCCAA	AATGTCGTAA	CAACTCCGCC	CCATTGACGC	AAATGGGCGG	TAGGCGTGTA	900
CGGTGGGAGG	TCTATATAAG	CAGAGCTGGG	TACGTGAACC	GTCAGATCGC	CTGGAGACGC	960
CATCACAGAT	CTCTCACTAT	GGATTTTCAG	GTGCAGATTA	TCAGCTTCCT	GCTAATCAGT	1020
GCTTCAGTCA	TAATGTCCAG	AGGACAAATT	GTTCTCTCCC	AGTCTCCAGC	AATCCTGTCT	1080
GCATCTCCAG	GGGAGAAGGT	CACAATGACT	TGCAGGGCCA	GCTCAAGTGT	AAGTTACATC	1140
CACTGGTTCC	AGCAGAAGCC	AGGATCCTCC	CCCAAACCTT	GGATTTATGC	CACATCCAAC	1200
CTGGCTTCTG	GAGTCCCTGT	TCGCTTCAGT	GGCAGTGGGT	CTGGGACTTC	TTACTCTCTC	1260
ACAATCAGCA	GAGTGGAGGC	TGAAGATGCT	GCCACTTATT	ACTGCCAGCA	GTGGACTAGT	1320
AACCCACCCA	CGTTCGGAGG	GGGGACCAAG	CTGGAAATCA	AACGTACGGT	GGCTGCACCA	1380
TCTGCTCTTA	CTTTCCCGCC	ATCTGATGAG	CAGTTGAAAT	CTGGAACTGC	CTCTGTTGTG	1440
TGCCTGTGTA	ATAACTTCTA	TCCCAGAGAG	GCCAAAGTAC	AGTGGAAGGT	GGATAACGCC	1500
CTCCAATCGG	GTAACTCCCA	GGAGAGTGTC	ACAGAGCAGG	ACAGCAAGGA	CAGCACCTAC	1560
AGCCTCAGCA	GCACCCTGAC	GCTGAGCAAA	GCAGACTACG	AGAAACACAA	AGTCTACGCC	1620
TGCGAAGTCA	CCCATCAGGG	CCTGAGCTCG	CCCGTCACAA	AGAGCTTCAA	CAGGGGAGAG	1680
TGTTGAATTC	AGATCCGTTA	ACGGTTACCA	ACTACCTAGA	CTGGATTTCGT	GACAACATGC	1740
GGCCGTGATA	TCTACGTATG	ATCAGCCTCG	ACTGTGCCTT	CTAGTTGCCA	GCCATCTGTT	1800
GTTTGCCCTT	CCCCCGTGCC	TTCTTGGACC	CTGGAAGGTG	CCACTCCCAC	TGTCCTTTCC	1860
TAATAAAATG	AGGAAATTGC	ATCGCATTGT	CTGAGTAGGT	GTCAATTCTAT	TCTGGGGGGT	1920
GGGGTGGGGC	AGGACAGCAA	GGGGGAGGAT	TGGGAAGACA	ATAGCAGGCA	TGCTGGGGAT	1980
GCGGTGGGCT	CTATGGAACC	AGCTGGGGCT	CGACAGCTAT	GCCAAGTACG	CCCCCTATTG	2040
ACGTCAATGA	CGGTAAATGG	CCCGCCTGGC	ATTATGCCCA	GTACATGACC	TTATGGGACT	2100
TTCTTACTTG	GCAGTACATC	TACGTATTAG	TCATCGCTAT	TACCATGGTG	ATGCGGTTTT	2160
GGCAGTACAT	CAATGGGCGT	GGATAGCGGT	TTGACTCACG	GGGATTTCCA	AGTCTCCACC	2220
CCATTGACGT	CAATGGGAGT	TTGTTTTGGC	ACCAAAATCA	ACGGGACTTT	CCAAAATGTC	2280
GTAACAACCTC	CGCCCCATTG	ACGCAAATGG	GCGGTAGGCG	TGTACGGTGG	GAGGTCTATA	2340
TAAGCAGAGC	TGGGTACGTC	CTCACATTCA	GTGATCAGCA	CTGAACACAG	ACCCGTCGAC	2400
ATGGGTTGGA	GCCTCATCTT	GCTCTTCCTT	GTCGCTGTTG	CTACGCGTGT	CCTGTCCCAG	2460
GTACAACCTG	AGCAGCCTGG	GGCTGAGCTG	GTGAAGCCTG	GGGCCTCAGT	GAAGATGTCC	2520
TGCAAGGCTT	CTGGCTACAC	ATTTACCAGT	TACAATATGC	ACTGGGTAAA	ACAGACACCT	2580

GGTCGGGGCC	TGGAATGGAT	TGGAGCTATT	TATCCCGGAA	ATGGTGATAC	TTCCTACAAT	2640
CAGAAGTTCA	AAGGCAAGGC	CACATTGACT	GCAGACAAAT	CCTCCAGCAC	AGCCTACATG	2700
CAGCTCAGCA	GCCTGACATC	TGAGGACTCT	GCGGTCTATT	ACTGTGCAAG	ATCGACTTAC	2760
TACGGCGGTG	ACTGGTACTT	CAATGTCTGG	GGCGCAGGGA	CCACGGTCAC	CGTCTCTGCA	2820
GCTAGACCA	AGGGCCCATC	GGTCTTCCCC	CTGGCACCCCT	CCTCCAAGAG	CACCTCTGGG	2880
GGCACAGCGG	CCCTGGGCTG	CCTGGTCAAG	GACTACTTCC	CCGAACCGGT	GACGGTGTCTG	2940
TGGAATCAG	GCGCCCTGAC	CAGCGGCGTG	CACACCTTCC	CGGCTGTCTT	ACAGTCCTCA	3000
GGACTCTACT	CCCTCAGCAG	CGTGGTGACC	GTGCCCTCCA	GCAGCTTGGG	CACCCAGACC	3060
TACATCTGCA	ACGTGAATCA	CAAGCCCAGC	AACACCAAGG	TGGACAAGAA	AGCAGAGCCC	3120
AAATCTTGTG	ACAAAACCTCA	CACATGCCCA	CCGTGCCCAG	CACCTGAACT	CCTGGGGGGA	3180
CCGTCAGTCT	TCCTCTTCCC	CCCAAAAACCC	AAGGACACCC	TCATGATCTC	CCGGACCCCT	3240
GAGGTCACAT	GCGTGGTGGT	GGACGTGAGC	CACGAAGACC	CTGAGGTCAA	GTTCAACTGG	3300
TACGTGGACG	GCGTGGAGGT	GCATAATGCC	AAGACAAAGC	CGCGGGAGGA	GCAGTACAAC	3360
AGCACGTACC	GTGTGGTCAG	CGTCCCTCACC	GTCCTGCACC	AGGACTGGCT	GAATGGCAAG	3420
GAGTACAAGT	GCAAGGTCTC	CAACAAAGCC	CTCCCAGCCC	CCATCGAGAA	AACCATCTCC	3480
AAAGCCAAAG	GGCAGCCCCG	AGAACCACAG	GTGTACACCC	TGCCCCCATC	CCGGGATGAG	3540
CTGACCAAGA	ACCAGGTCAG	CCTGACCTGC	CTGGTCAAAG	GCTTCTATCC	CAGCGACATC	3600
GCCGTGGAGT	GGGAGAGCAA	TGGGCAGCCG	GAGAACAAC	ACAAGACCAC	GCCTCCCGTG	3660
CTGGACTCCG	ACGGCTCCTT	CTTCCCTCTAC	AGCAAGCTCA	CCGTGGACAA	GAGCAGGTGG	3720
CAGCAGGGGA	ACGTCTTCTC	ATGCTCCGTG	ATGCATGAGG	CTCTGCACAA	CCACTACACG	3780
CAGAAGAGCC	TCTCCCTGTC	TCCGGGTAA	TGAGGATCCG	TTAACGGTTA	CCAACCTACCT	3840
AGACTGGATT	CGTGACAACA	TGCGGCCGTG	ATATCTACGT	ATGATCAGCC	TCGACTGTGC	3900
CTTCTAGTTG	CCAGCCATCT	GTTGTTTGCC	CCTCCCCCGT	GCCTTCCTTG	ACCCTGGAAG	3960
GTGCCACTCC	CACCTGCTCT	TCCTAATAAA	ATGAGGAAAT	TGCATCGCAT	TGTCTGAGTA	4020
GGTGTCAATC	TATTCTGGGG	GGTGGGGTGG	GGCAGGACAG	CAAGGGGGAG	GATTGGGAAG	4080
ACAATAGCAG	GCATGCTGGG	GATGCGGTGG	GCTCTATGGA	ACCAGCTGGG	GCTCGACAGC	4140
GCTGGATCTC	CCGATCCCCA	GCTTTGCTTC	TCAATTTCTT	ATTTGCATAA	TGAGAAAAAA	4200
AGGAAAATTA	ATTTTAACAC	CAATTCAGTA	GTTGATTGAG	CAAATGCGTT	GCCAAAAAGG	4260
ATGCTTTAGA	GACAGTGTTT	TCTGCACAGA	TAAGGACAAA	CATTATTTCAG	AGGGAGTACC	4320
CAGAGCTGAG	ACTCCTAAGC	CAGTGAGTGG	CACAGCATTC	TAGGGAGAAA	TATGCTTGTC	4380
ATCACCGAAG	CCTGATTCCG	TAGAGCCACA	CCTTGGTAA	GGCCAATCTG	CTCACACAGG	4440
ATAGAGAGGG	CAGGAGCCAG	GGCAGAGCAT	ATAAGGTGAG	GTAGGATCAG	TTGCTCCTCA	4500
CATTTGCTTC	TGACATAGTT	GTGTTGGGAG	CCTTGATAGC	TTGGACAGCT	CAGGGCTGCG	4560
ATTTTCGCGC	AAACTTGACG	GCAATCCTAG	CGTGAAGGCT	GGTAGGATTT	TATCCCCGCT	4620
GCCATCATGG	TTCGACCATT	GAATGTCATC	GTCGCCGTGT	CCCAAAATAT	GGGGATTGGC	4680
AAGAACGGAG	ACCTACCCTG	GCCTCCGCTC	AGGAACGAGT	TCAAGTACTT	CCAAAGAATG	4740
ACCACAACCT	CTTCAGTGGA	AGGTAAACAG	AATCTGGTGA	TTATGGGTAG	GAAAACCTGG	4800
TTCTCCATT	CTGAGAAGAA	TCGACCTTTA	AAGGACAGAA	TTAATATAGT	TCTCAGTAGA	4860
GAATCAAA	AACCACCACG	AGGAGCTCAT	TTTCTTGCCA	AAAAGTTTGA	TGATGCCTTA	4920
AGACTTATTG	AACAACCGGA	ATTGGCAAGT	AAAAGTAGACA	TGGTTTGGAT	AGTCGGAGGC	4980
AGTTCTGT	ACCAGGAAGC	CATGAATCAA	CCAGGCCACC	TTAGACTCTT	TGTGACAAGG	5040
ATCATGCAGG	AATTTGAAAG	TGACACGTTT	TTCCAGAAA	TTGATTTGGG	GAAATATAAA	5100
CTTCTCCCAG	AATACCCAGG	CGTCCCTCT	GAGGTCCAGG	AGGAAAAAGG	CATCAAGTAT	5160
AAGTTTGAAG	TCTACGAGAA	GAAAGACTAA	CAGGAAGATG	CTTCAAGTT	CTCTGCTCCC	5220
CTCCTAAAGC	TATGCATTTT	TATAAGACCA	TGGGACTTTT	GCTGGCTTTA	GATCAGCCTC	5280
GACTGTGCCT	TCTAGTTGCC	AGCCATCTGT	TGTTTGCCCC	TCCCCCGTGC	CTTCCCTTGAC	5340
CCTGGAAGGT	GCCACTCCCA	CTGTCCCTTC	CTAATAAAAT	GAGGAAATTG	CATCGCATTG	5400
TCTGAGTAGG	TGTCAATTCTA	TTCTGGGGGG	TGGGGTGGGG	CAGGACAGCA	AGGGGGAGGA	5460
TTGGGAAGAC	AATAGCAGGC	ATGCTGGGGG	TGCGGTGGGC	TCTATGGAAC	CAGCTGGGGC	5520
TCGAGCTACT	AGCTTTTGCTT	CTCAATTTCT	TATTTGCATA	ATGAGAAAAA	AAGGAAAATT	5580
AAATTTAACA	CCAATTCAGT	AGTTGATTGA	GCAATGCGT	TGCCAAAAAG	GATGCTTTAG	5640
AGACAGTGTT	CTCTGCACAG	ATAAGGACAA	ACATTATTCA	GAGGGAGTAC	CCAGAGCTGA	5700
GACTCCTAAG	CCAGTGAGTG	GCACAGCATT	CTAGGGAGAA	ATATGCTTGT	CATCACCGAA	5760
GCCTGATTCC	GTAGAGCCAC	ACCTTGGTAA	GGGCCAATCT	GCTCACACAG	GATAGAGAGG	5820
GCAGGAGCCA	GGGCAGAGCA	TATAAGGTGA	GGTAGGATCA	GTTGCTCCTC	ACATTTGCTT	5880
CTGACATAGT	TGTGTTGGGA	GCTTGGATCG	ATCCTCTATG	GTTGAACAAG	ATGGATTGCA	5940
CGCAGGTTCT	CCGGCCGCTT	GGGTGGAGAG	GCTATTCCGG	TATGACTGGG	CACAACAGAC	6000
AATCGGCTGC	TCTGATGCCG	CCGTGTTCCG	GCTGTCAGCG	CAGGGGCGCC	CGGTTCTTTT	6060

TGTCAAGACC	GACCTGTCCG	GTGCCCTGAA	TGAACTGCAG	GACGAGGCAG	CGCGGCTATC	6120
GTGGCTGGCC	ACGACGGGCG	TTCCTTGCGC	AGCTGTGCTC	GACGTTGTCA	CTGAAGCGGG	6180
AAGGGACTGG	CTGCTATTGG	GCGAAGTGCC	GGGGCAGGAT	CTCCTGTCAT	CTCACCTTGC	6240
TCCTGCCGAG	AAAGTATCCA	TCATGGCTGA	TGCAATGCGG	CGGCTGCATA	CGCTTGATCC	6300
GGCTACCTGC	CCATTTCGACC	ACCAAGCGAA	ACATCGCATC	GAGCGAGCAC	GTACTCGGAT	6360
GGAAGCCGGT	CTTGTCGATC	AGGATGATCT	GGACGAAGAG	CATCAGGGGC	TCGCGCCAGC	6420
CGAACTGTTC	GCCAGGCTCA	AGGCGCGCAT	GCCCCACGGC	GAGGATCTCG	TCGTGACCCA	6480
TGGCGATGCC	TGCTTGCCGA	ATATCATGGT	GGAAAATGGC	CGCTTTTCTG	GATTTCATCGA	6540
CTGTGGCCGG	CTGGGTGTGG	CGGACCGCTA	TCAGGACATA	GCGTTGGCTA	CCCGTGATAT	6600
TGCTGAAGAG	CTTGGCGGCG	AATGGGCTGA	CCGCTTCCTC	GTGCTTTACG	GTATCGCCGC	6660
TCCCCGATTCG	CAGCGCATCG	CCTTCTATCG	CCTTCTTGAC	GAGTTCTTCT	GAGCGGGACT	6720
CTGGGGTTCG	AAATGACCGA	CCAAGCGACG	CCCAACCTGC	CATCACGAGA	TTTCGATTCC	6780
ACCGCCGCCT	TCTATGAAAG	GTTGGGCTTC	GGAATCGTTT	TCCGGGACGC	CGGCTGGATG	6840
ATCCTCCAGC	GCGGGGATCT	CATGCTGGAG	TTCTTCGCCC	ACCCCAACTT	GTTTATTGCA	6900
GCTTATAATG	GTTACAAATA	AAGCAATAGC	ATCACAAATT	TCACAAATAA	AGCATTTTTT	6960
TCACTGCATT	CTAGTTGTGG	TTTGTCCAAA	CTCATCAATC	TATCTTATCA	TGTCTGGATC	7020
GCGGCCGCGA	TCCCGTCGAG	AGCTTGGCGT	AATCATGGTC	ATAGCTGTTT	CCTGTGTGAA	7080
ATTGTTATCC	GCTCACAAAT	CCACACAACA	TACGAGCCGG	AAGCATAAAG	TGTAAAGCCT	7140
GGGGTGCCTA	ATGAGTGAGC	TAACTCACAT	TAATTGCGTT	GCGCTCACTG	CCCGCTTTCC	7200
AGTCGGGAAA	CCTGTCGTGC	CAGCTGCATT	AATGAATCGG	CCAACGCGCG	GGGAGAGGCG	7260
GTTTGCGTAT	TGGGCGCTCT	TCCGCTTCCT	CGCTCACTGA	CTCGCTGCGC	TCGGTCGTTC	7320
GGCTGCGGCG	AGCGGTATCA	GCTCACTCAA	AGGCGGTAAT	ACGGTTATCC	ACAGAATCAG	7380
GGGATAACGC	AGGAAAGAAC	ATGTGAGCAA	AAGGCCAGCA	AAAGGCCAGG	AACCGTAAAA	7440
AGGCCGCGTT	GCTGGCGTTT	TTCCATAGGC	TCCGCCCCC	TGACGAGCAT	CACAAAAATC	7500
GACGCTCAAG	TCAGAGGTGG	CGAAACCCGA	CAGGACTATA	AAGATACCAG	GCGTTTCCCC	7560
CTGGAAGCTC	CCTCGTGCGC	TCTCCTGTTT	CGACCCTGCC	GCTTACCGGA	TACCTGTCCG	7620
CCTTTCTCCC	TTCGGGAAGC	GTGGCGCTTT	CTCAATGCTC	ACGCTGTAGG	TATCTCAGTT	7680
CGGTGTAGGT	CGTTCGCTCC	AAGCTGGGCT	GTGTGCACGA	ACCCCCCGTT	CAGCCCGACC	7740
GCTGCGCCTT	ATCCGGTAAC	TATCGTCTTG	AGTCCAACCC	GGTAAGACAC	GACTTATCGC	7800
CACTGGCAGC	AGCCACTGGT	AACAGGATTA	GCAGACGAG	GTATGTAGGC	GGTGCTACAG	7860
AGTTCTTGAA	GTTGGTGGCT	AACTACGGCT	ACACTAGAAG	GACAGTATTT	GGTATCTGCG	7920
CTCTGCTGAA	GTCAGTTACC	TTTCGAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAACAAA	7980
CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAG	8040
GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT	8100
CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA	8160
ATTAAAAATG	AAGTTTAAAA	TCAATCTAAA	GTATATATGA	GTAAACTTGG	TCTGACAGTT	8220
ACCAATGCTT	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	TCTATTTCTG	TCATCCATAG	8280
TTGCTGACT	CCCCGTCGTG	TAGATAACTA	CGATACGGGA	GGGCTTACCA	TCTGGCCCCA	8340
GTGCTGCAAT	GATACCGCGA	GACCCACGCT	CACCGGCTCC	AGATTTATCA	GCAATAAACC	8400
AGCCAGCCGG	AAGGGCCGAG	CGCAGAAGTG	GTCCTGCAAC	TTTATCCGCC	TCCATCCAGT	8460
CTATTAATTG	TTGCCGGGAA	GCTAGAGTAA	GTAGTTCGCC	AGTTAATAGT	TTGCGCAACG	8520
TTGTTGCCAT	TGCTACAGGC	ATCGTGGTGT	CACGCTCGTC	GTTTGGTATG	GCTTCATTCA	8580
GCTCCGGTTC	CCAACGATCA	AGGCGAGTTA	CATGATCCCC	CATGTTGTGC	AAAAAAGCGG	8640
TTAGCTCCTT	CGGTCCTCCG	ATCGTTGTCA	GAAGTAAGTT	GGCCGCAGTG	TTATCACTCA	8700
TGTTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTCAATG	ATCCGTAAAG	TGCTTTTCTG	8760
TGACTGGTGA	GTACTCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT	8820
CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA	8880
TCATTGGAAA	ACGTTCTTCG	GGGCGAAAAA	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA	8940
GTTTCGATGA	ACCCACTCGT	GCACCCAACT	GATCTTCAGC	ATCTTTTACT	TTCACCAGCG	9000
TTTCTGGGTG	AGCAAAAACA	GGAAGGCAAA	ATGCCGCAAA	AAAGGGAATA	AGGGCGACAC	9060
GGAAATGTTG	AATACTCATA	CTCTTCCTTT	TTCAATATTA	TTGAAGCATT	TATCAGGGTT	9120
ATTGTCTCAT	GAGCGGATAC	ATATTTGAAT	GTATTTAGAA	AAATAAACAA	ATAGGGGTTT	9180
CGCGCACATT	TCCCCGAAAA	GTGCCACCT				9209

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATCACAGATC TCTCACCATG GATTTTCAGG TGCAGATTAT CAGCTTC

47

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "primer"

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TGCAGCATCC GTACGTTTGA TTTCCAGCTT

30

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 384 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..384

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATG GAT TTT CAG GTG CAG ATT ATC AGC TTC CTG CTA ATC AGT GCT TCA  
Met Asp Phe Gln Val Gln Ile Ile Ser Phe Leu Leu Ile Ser Ala Ser  
1 5 10 15

48

GTC ATA ATG TCC AGA GGG CAA ATT GTT CTC TCC CAG TCT CCA GCA ATC  
Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile  
20 25 30

96

CTG TCT GCA TCT CCA GGG GAG AAG GTC ACA ATG ACT TGC AGG GCC AGC	144
Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser	
35 40 45	
TCA AGT GTA AGT TAC ATC CAC TGG TTC CAG CAG AAG CCA GGA TCC TCC	192
Ser Ser Val Ser Tyr Ile His Trp Phe Gln Gln Lys Pro Gly Ser Ser	
50 55 60	
CCC AAA CCC TGG ATT TAT GCC ACA TCC AAC CTG GCT TCT GGA GTC CCT	240
Pro Lys Pro Trp Ile Tyr Ala Thr Ser Asn Leu Ala Ser Gly Val Pro	
65 70 75 80	
GTT CGC TTC AGT GGC AGT GGG TCT GGG ACT TCT TAC TCT CTC ACA ATC	288
Val Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile	
85 90 95	
AGC AGA GTG GAG GCT GAA GAT GCT GCC ACT TAT TAC TGC CAG CAG TGG	336
Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp	
100 105 110	
ACT AGT AAC CCA CCC ACG TTC GGA GGG GGG ACC AAG CTG GAA ATC AAA	384
Thr Ser Asn Pro Pro Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys	
115 120 125	

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 128 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Asp Phe Gln Val Gln Ile Ile Ser Phe Leu Leu Ile Ser Ala Ser	
1 5 10 15	
Val Ile Met Ser Arg Gly Gln Ile Val Leu Ser Gln Ser Pro Ala Ile	
20 25 30	
Leu Ser Ala Ser Pro Gly Glu Lys Val Thr Met Thr Cys Arg Ala Ser	
35 40 45	
Ser Ser Val Ser Tyr Ile His Trp Phe Gln Gln Lys Pro Gly Ser Ser	
50 55 60	
Pro Lys Pro Trp Ile Tyr Ala Thr Ser Asn Leu Ala Ser Gly Val Pro	
65 70 75 80	
Val Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile	
85 90 95	
Ser Arg Val Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp	
100 105 110	

Thr Ser Asn Pro Pro Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys  
 115 120 125

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCGGCTCCCA CGCGTGTCTT GTCCAG

27

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
 (A) DESCRIPTION: /desc = "primer"

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGSTGTTGTG CTAGCTGMRG AGACRGTA

29

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 420 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..420

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATG GGT TGG AGC CTC ATC TTG CTC TTC CTT GTC GCT GTT GCT ACG CGT	48
Met Gly Trp Ser Leu Ile Leu Leu Phe Leu Val Ala Val Ala Thr Arg	
1 5 10 15	
GTC CTG TCC CAG GTA CAA CTG CAG CAG CCT GGG GCT GAG CTG GTG AAG	96
Val Leu Ser Gln Val Gln Leu Gln Gln Pro Gly Ala Glu Leu Val Lys	
20 25 30	
CCT GGG GCC TCA GTG AAG ATG TCC TGC AAG GCT TCT GGC TAC ACA TTT	144
Pro Gly Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe	
35 40 45	
ACC AGT TAC AAT ATG CAC TGG GTA AAA CAG ACA CCT GGT CGG GGC CTG	192
Thr Ser Tyr Asn Met His Trp Val Lys Gln Thr Pro Gly Arg Gly Leu	
50 55 60	
GAA TGG ATT GGA GCT ATT TAT CCC GGA AAT GGT GAT ACT TCC TAC AAT	240
Glu Trp Ile Gly Ala Ile Tyr Pro Gly Asn Gly Asp Thr Ser Tyr Asn	
65 70 75 80	
CAG AAG TTC AAA GGC AAG GCC ACA TTG ACT GCA GAC AAA TCC TCC AGC	288
Gln Lys Phe Lys Gly Lys Ala Thr Leu Thr Ala Asp Lys Ser Ser Ser	
85 90 95	
ACA GCC TAC ATG CAG CTC AGC AGC CTG ACA TCT GAG GAC TCT GCG GTC	336
Thr Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val	
100 105 110	
TAT TAC TGT GCA AGA TCG ACT TAC TAC GGC GGT GAC TGG TAC TTC AAT	384
Tyr Tyr Cys Ala Arg Ser Thr Tyr Tyr Gly Gly Asp Trp Tyr Phe Asn	
115 120 125	
GTC TGG GGC GCA GGG ACC ACG GTC ACC GTC TCT GCA	420
Val Trp Gly Ala Gly Thr Thr Val Thr Val Ser Ala	
130 135 140	

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 140 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Gly Trp Ser Leu Ile Leu Leu Phe Leu Val Ala Val Ala Thr Arg	
1 5 10 15	
Val Leu Ser Gln Val Gln Leu Gln Gln Pro Gly Ala Glu Leu Val Lys	
20 25 30	
Pro Gly Ala Ser Val Lys Met Ser Cys Lys Ala Ser Gly Tyr Thr Phe	
35 40 45	

Thr	Ser	Tyr	Asn	Met	His	Trp	Val	Lys	Gln	Thr	Pro	Gly	Arg	Gly	Leu
	50					55					60				
Glu	Trp	Ile	Gly	Ala	Ile	Tyr	Pro	Gly	Asn	Gly	Asp	Thr	Ser	Tyr	Asn
65					70					75					80
Gln	Lys	Phe	Lys	Gly	Lys	Ala	Thr	Leu	Thr	Ala	Asp	Lys	Ser	Ser	Ser
				85					90					95	
Thr	Ala	Tyr	Met	Gln	Leu	Ser	Ser	Leu	Thr	Ser	Glu	Asp	Ser	Ala	Val
			100					105					110		
Tyr	Tyr	Cys	Ala	Arg	Ser	Thr	Tyr	Tyr	Gly	Gly	Asp	Trp	Tyr	Phe	Asn
		115					120					125			
Val	Trp	Gly	Ala	Gly	Thr	Thr	Val	Thr	Val	Ser	Ala				
	130					135					140				